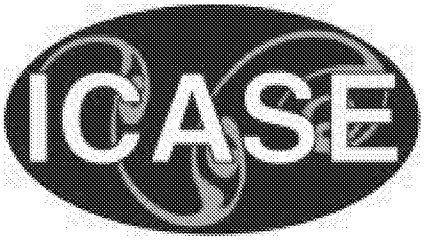
NASA/CR-2001-211239
ICASE Report No. 2001-35



Random Field Solutions Including Boundary Condition Uncertainty for the Steady-state Generalized Burgers Equation

Luc Huyse
ICASE, Hampton, Virginia

Robert W. Walters
Virginia Polytechnic Institute and State University, Blacksburg, Virginia

October 2001

The NASA STI Program Office . . . in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA Scientific and Technical Information (STI) Program Office plays a key part in helping NASA maintain this important role.

The NASA STI Program Office is operated by Langley Research Center, the lead center for NASA's scientific and technical information. The NASA STI Program Office provides access to the NASA STI Database, the largest collection of aeronautical and space science STI in the world. The Program Office is also NASA's institutional mechanism for disseminating the results of its research and development activities. These results are published by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA's counterpart of peer-reviewed formal professional papers, but having less stringent limitations on manuscript length and extent of graphic presentations.

- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

- CONFERENCE PUBLICATIONS. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or cosponsored by NASA.

- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.

- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services that complement the STI Program Office's diverse offerings include creating custom thesauri, building customized data bases, organizing and publishing research results . . . even providing videos.

For more information about the NASA STI Program Office, see the following:

- Access the NASA STI Program Home Page at ***http://www.sti.nasa.gov***

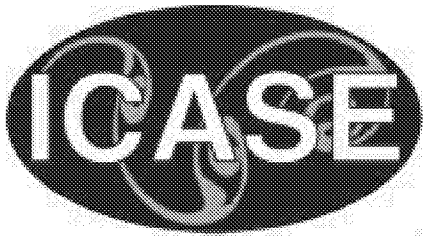
- Email your question via the Internet to help@sti.nasa.gov

- Fax your question to the NASA STI Help Desk at (301) 621-0134

- Telephone the NASA STI Help Desk at (301) 621-0390

- Write to:
  NASA STI Help Desk
  NASA Center for AeroSpace Information
  7121 Standard Drive
  Hanover, MD 21076-1320

NASA/CR-2001-211239
ICASE Report No. 2001-35



Random Field Solutions Including Boundary Condition Uncertainty for the Steady-state Generalized Burgers Equation

Luc Huyse
ICASE, Hampton, Virginia

Robert W. Walters
Virginia Polytechnic Institute and State University, Blacksburg, Virginia

ICASE
NASA Langley Research Center
Hampton, Virginia

Operated by Universities Space Research Association

National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23681-2199

Prepared for Langley Research Center
under Contract NAS1-97046

October 2001

Available from the following:

NASA Center for AeroSpace Information (CASI)
7121 Standard Drive
Hanover, MD 21076-1320
(301) 621-0390

National Technical Information Service (NTIS)
5285 Port Royal Road
Springfield, VA 22161-2171
(703) 487-4650

RANDOM FIELD SOLUTIONS INCLUDING BOUNDARY CONDITION UNCERTAINTY FOR THE STEADY-STATE GENERALIZED BURGERS EQUATION

LUC HUYSE* AND ROBERT W. WALTERS†

Abstract. CFD results are subject to considerable uncertainty associated with the operating conditions. Even when the operational uncertainty is omitted under very controlled circumstances during wind tunnel experiments, substantial disagreement between experimental and CFD results persists. This discrepancy must be attributed to model uncertainty. This report discusses the various sources of uncertainty. The need for advanced uncertainty modeling is illustrated by means of a computationally inexpensive 1-D Burgers equation model. We specifically address the uncertainty due to missing variables (inexact or incomplete differential equations). To this extent a random field model is used for the viscosity and the fundamental differences between the solutions of the stochastic differential equations and a simple random variable model is highlighted. The Burgers equation theoretically needs to be integrated over an infinite domain. In a deterministic approach, the integration domain is cut off at some far field boundary. This truncation effectively ignores all variability outside this far field boundary. We present a practical treatment for the uncertainty on the boundary conditions. The results indicate that ignoring the boundary condition uncertainty dramatically underestimates the variance of the velocity $u(x)$ in the interior of the domain.

Key words. model uncertainty, random field, uncertain boundary conditions

Subject classification. Applied and Numerical Mathematics

1. Introduction & Motivation. Both computational results and experimental measurements in aerodynamic applications are subject to considerable uncertainty as illustrated by the scatter in the solutions submitted by various researchers for a recent AIAA drag prediction workshop [1]. For these reasons, the current certification process relies heavily on full-scale flight testing. Dramatic savings would be achieved if the need for expensive full-scale flight tests could be reduced through improved reliability (or dependability) of CFD results.

The interest in uncertainty modeling within the CFD community has grown considerably in recent years. In a deterministic approach, the performance of a design is typically assessed for a limited number of design or operating conditions. Uncertainty associated with the operating conditions (variable payload, atmospheric conditions) and fluctuations in the geometry and smoothness of the skin (manufacturing uncertainty) may have substantial impact on the results. Techniques for propagating these uncertainties require additional computational effort but are well established, an example application is described in [30]. Exact and approximate uncertainty assessment methods are applied to an airfoil optimization problem in [17], where a dramatic improvement of the robustness of the design was achieved.

However, even if these outside sources of "operational" uncertainty are omitted under very controlled circumstances during wind tunnel experiments, substantial disagreement between experimental and CFD results persists [1]. Assuming that the CFD model and the experiment simulate the same thing, this discrepancy must be attributed to model uncertainty. In this paper we will address the issue of modeling uncertainty for a computationally inexpensive 1-D Burgers equation model.

*ICASE, Mail Stop 132C, NASA Langley Research Center, Hampton, VA 23681-2199 (email: l.huyse@larc.nasa.gov). This research was supported by the National Aeronautics and Space Administration under NASA Contract No. NAS1-97046 while the author was in residence at ICASE, NASA Langley Research Center, Hampton, VA 23681-2199.

†Virginia Polytechnic Institute and State University, Department of Aerospace and Ocean Engineering, Blacksburg, VA 24061 (email: walters@aoe.vt.edu). This research was supported by the National Aeronautics and Space Administration under NASA Contract No. NAS1-97046 while the author was in residence at ICASE, NASA Langley Research Center, Hampton, VA 23681-2199.

2. Outline of Paper. After an initial discussion of the different sources of uncertainty and quantitative uncertainty modeling techniques in Sections 3 and 4, the Burgers equation is introduced in Section 5. Results for a random variable model of the viscosity are presented in Section 6. A random variable model of the viscosity assumes that the viscosity remains constant throughout the entire domain. However, the temperature dependence of the viscosity introduces additional uncertainty in the viscosity which is best captured using a random field model. This model uncertainty effectively transforms the Burgers equation into a stochastic differential equation. Second-order random field models are discussed in Sections 7. Two discretization techniques, midpoint and locally averaged discretizations, are presented in Sections 8 and 9. Results for the random field model are given in Section 10. The boundary conditions for the Burgers equation are specified at infinity. In numerical solution methods, the integration domain is limited and cut off at the "far field" boundary. The additional uncertainty due to this boundary condition treatment is assessed in Section 11. Final conclusions are drawn in Section 12.

3. Sources of Uncertainty. At this point in the discussion it is appropriate to spend some time on nomenclature, especially the difference between "error", "scatter" and "uncertainty". Unfortunately, these terms have been used to denote somewhat different things in different fields of application. Historically, the CFD community has sometimes used the term "uncertainty" to denote "error" in relation to grid convergence studies [32]. This confounds the issue. We adopt the following definitions, which are commonly used by statisticians:

- Error is a deterministic concept and is defined as the difference between the true or exact answer to a problem and the answer, computed or measured using a faulty or simplified theory.
- Scatter measures the range or spread of the data, but gives no information about the potential bias due to systematic measurement error or due to missing terms in the CFD code.
- Uncertainty indicates that the result can be only known with a limited amount of confidence for a given level of precision. This uncertainty is an inherent property of the measurement technique or model description and is due to lack of knowledge.

The distinction between the deterministic and stochastic nature is apparent in the AIAA definitions as well [2]:

- Error is a recognizable deficiency in any phase or activity of modeling and simulation that is not due to lack of knowledge.
- Uncertainty is a potential deficiency in any phase or activity of modeling and simulation that is due to lack of knowledge.

Since error is a recognizable deficiency, all errors are – in principle at least – correctable and therefore deterministic. Since uncertainty is caused by a fundamental lack of knowledge, it cannot be eliminated. If a higher confidence level (or level of credibility) of the prediction is required, the result can only be given with less precision. Uncertainty forces a fundamental trade-off between confidence and precision.

Two sources of uncertainty can be distinguished [6],[28]: inherent (aleatoric) and model (or epistemic) uncertainty.

- Inherent Uncertainty: this applies to phenomena which we accept to be intrinsically variable in our model. A descriptive, often statistical, formulation is used for this uncertainty and no attempt is made to explain any of this variability. In a mathematical model, these uncertainties are typically described using (joint) probability density functions. Techniques for the efficient propagation of these distributions through a model will be briefly discussed in this report.
- Model Uncertainty: three types can be distinguished here. The first two types can be caused by either "errors of ignorance" or "errors of simplification". Note that use of the word "error" means that these uncertainties originate in either an acknowledged deliberate simplification or an acknowledged lack of understanding. However, since the exact solution is unknown is this case, the errors can not be corrected and we have no

alternative but to treat them in a non-deterministic manner as uncertainties.

1. missing variables: a typical example is the exclusion of coupling effects from the analysis. For instance the viscosity depends on the temperature, but we may decide not to do a full coupled aero-thermal analysis and ignore the effect of the temperature on the viscosity. Ignoring structural deformations of the wing in an aerodynamic analysis is another example thereof.
2. inexact functional form: quite often we use linear equations whereas the original physics model contains non-linear terms. The linearization introduces an acknowledged error, however, we may be (practically) unable to solve the original non-linear equations.
3. parameter (or statistical) uncertainty: provided the n parameters in the model exist and are unique, their values can be determined from n experiments if both the model and experimental data are exact. However, in the presence of incomplete information (measurement error, model error), no amount of data can provide exact solutions for them and the true values of these parameters will remain uncertain.

4. Quantitative Uncertainty Modeling. Uncertainty quantification can be accomplished using either probabilistic or deterministic methods [28]. Examples of non-probabilistic, sometimes also referred to as possibilistic methods, are interval analysis [15], Dempster-Shafer theory [36], convex modeling [5], and fuzzy computation methods [31]. Broadly speaking these methods address the possibility of an event within the bounds imposed on the uncertain parameters. They do, however, not give any information regarding the likelihood of this event taking place.

Probabilistic methods, on the other hand, provide ample quantitative information regarding the likelihood of an event taking place. For varying degrees of computational effort, one can compute statistics, and the confidence intervals thereof, or even complete probability density functions (PDF) for any particular quantity of interest. It is however, important to realize that the model outcomes can be very sensitive to the accuracy of the provided probabilistic input [11]. It is important to recognize that not just marginal PDFs are required but that an accurate estimation of the correlation between random variables is paramount [41].

Most specialists agree that perhaps the most promising applications for the non-probabilistic methods will occur when reliable data are unavailable or hard to get. Non-probabilistic methods are mostly used in absence of reliable and accurate data. There is little or no debate that, when sufficient data exist, the probabilistic model is superior. Probabilistic computations give much more comprehensive knowledge about the state of a particular system. The debate between the believers in probabilistic and possibilistic methods will probably (or is it: possibly?) continue for considerable time. In short, the end-user should never forget that no matter what model (probabilistic or possibilistic) is used, the computed outcomes are conditional upon the model assumptions made [12]. Uncertainty associated with the models should not be taken as an excuse for sloppiness! No matter what quantitative uncertainty analysis model is used, the validity of the results depends on the validity of the model.

Recently, a mathematical framework which unifies the probabilistic and possibilistic algorithms has been suggested [19]; it consists of a general constrained optimization problem:

$$\min\{s(\mathbf{x})|g(\mathbf{x}) = 0\} \tag{4.1}$$

where $g(\mathbf{x})$ represents the limit state function, *i.e.* the zero safety margin, as a function of the uncertain parameters $\mathbf{x}$ and $s(\mathbf{x})$ is some function which characterizes the method. For instance, when $s(\mathbf{x})$ is set equal to the reliability index β, the solution to Eq. (4.1) reverts to the well-known first-order reliability method (FORM) [24].

In this paper, only probabilistic analysis methods are used. Based on the previous discussion of uncertainty types, the following 3 issues need to be addressed in any quantitative uncertainty assessment:

1. characterize uncertainty associated with the system parameters and outside environment
2. propagate these uncertainties through the (potentially large) system

3. incorporate the model uncertainty -lack of data or lack of knowledge about the system behavior- into the overall assessment

4.1. Characterizing aleatory uncertainty. A complete statistical description requires the identification of all marginal and joint probability density functions of each random variable in the problem [13]. In practice, the required amount of data simply will not be available or the data structure will be so complex that only an approximate description is mathematically tractable. One form of approximate modeling is the use of conditional distributions, where the less important random variables are modeled conditional upon the values of the more important random variables. Whether a particular random variable is important depends on both its own variability and on the sensitivity derivative. Alternatively a second-order description can be used, which consists of the marginal PDF for each of the variables and the correlation coefficients between them. The correlations can be hard to estimate accurately but the importance thereof can not be understated and is clearly illustrated, for instance, by the expression for the first-order estimate of the variance of $z = f(x, y)$

$$\mathrm{Var}(z) = \mathrm{Var}(x)\left(\frac{\partial f}{\partial x}\right)^2 + \mathrm{Var}(y)\left(\frac{\partial f}{\partial y}\right)^2 + 2\,\mathrm{Covar}(x,y)\left(\frac{\partial f}{\partial x}\right)\left(\frac{\partial f}{\partial y}\right) \tag{4.2}$$

where all derivatives are evaluated at the mean values $\mathbb{E}(x)$ and $\mathbb{E}(y)$ and where

$$\mathrm{Covar}(x,y) = \rho_{xy}\sqrt{\mathrm{Var}(x)\mathrm{Var}(y)} = \mathbb{E}\left[xy - \mathbb{E}(x)\mathbb{E}(y)\right] \tag{4.3}$$

is the covariance of x and y and ρ_{xy} is the correlation coefficient of x and y.

In the absence of a sufficient amount of data, expert opinion can be used to construct an approximate joint PDF. An overview of expert elicitation techniques is described in [26] and [14]. It should be stressed however, that significant statistical uncertainty is associated with such techniques (discussed in Section 4.3).

4.2. Propagating uncertainties. Consider the vector $x = (x_1, \ldots, x_n)$ and its joint PDF $f_X(x_1, \ldots, x_n)$. The relationship between the input variables x and the output variables $y = (y_1, \ldots, y_n)$ is defined by

$$y_i = g_i(x_1, \ldots, x_n) \text{ for } i = 1, \ldots, n \tag{4.4}$$

together with the inverse functions

$$x_i = g_i^{-1}(y_1, \ldots, y_n) \text{ for } i = 1, \ldots, n \tag{4.5}$$

The Jacobian J of the transformation is:

$$J = \begin{vmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial y_n}{\partial x_1} & \frac{\partial y_n}{\partial x_2} & \cdots & \frac{\partial y_n}{\partial x_n} \end{vmatrix} \tag{4.6}$$

The exact joint PDF for the model output $y = g(x)$ in terms of the y-variables is given by [16]:

$$\begin{aligned} f_Y(y_1, \ldots, y_n) &= |J|^{-1} f_X(x_1, \ldots, x_n) \\ &= |J|^{-1} f_X(g_1^{-1}(y_1, \ldots, y_n), \ldots, g_n^{-1}(y_1, \ldots, y_n)) \end{aligned} \tag{4.7}$$

where $f_\bullet(\bullet)$ is the PDF of the variable $\bullet$.

Various techniques of uncertainty propagation are briefly outlined in this section. Three levels of methods can be identified:

1. Simulation methods: In their most elementary form (brute force Monte Carlo) one samples repeatedly from the joint PDF for the uncertain variables and generates the histograms of the model response quantities of interest [25]. The method has the advantage that it is simple and universally applicable. The convergence does not directly depend on the number of random variables in the problem. In this form Monte Carlo methods always give the correct answer, but a prohibitively large number of simulations may be required to accurately estimate extreme responses, which have a small probability of occurrence. Quite often these are the ones of interest in structural engineering and for this reason alternative sampling techniques which lead to much faster convergence have been developed. An overview of such variance-reduction techniques is available in [13], [25] and [34]. Such techniques should be used carefully since importance sampling, e.g., can introduce severe bias if used inappropriately.

   However, the brute force Monte Carlo method requires a much smaller number of samples – typically around 100 or less – to accurately estimate the mean and standard deviation, which are the quantities of interest in computational fluid dynamics applications. Brute force Monte Carlo also sounds appealing since its implementation can be written as a wrapper around an existing, deterministic algorithm. This makes it easy to transform any deterministic algorithm into a probabilistic code. However, this approach is not very efficient, particularly when the deterministic code uses an iterative solution process. The combination of a limited amount of extra coding to make use of restarts and a smart sampling scheme may result in large pay-offs in execution time.
2. Second-moment analytic approximations: these methods are based on a Taylor series expansion around the mean value of the input variables. They work best when the input uncertainties are not too large (coefficient of variation less than 10 to 15%) and approximately Gaussian. The methods require accurate derivative information. They have proven highly accurate in structural mechanics applications [4].
3. Advanced analytic methods: These methods can be regarded as a generalized Monte Carlo method [34]. However, since the solution mechanics are so different, we treat them as a separate category. In such a method the stochastic output variables are written as a series expansion. This series expansion formulation readily accommodates advanced random field modeling of the uncertainties. The basic equations (ODE or PDE) are now solved stochastically and a large amount of coding is required. The trade-off is that an accurate solution can be obtained for moments or output PDFs which accounts for both the non-linearities in the differential equations and the non-Gaussian nature of some of the input variables. Polynomial Chaos [39] is one member of this class of methods.

4.3. Model uncertainty. Model uncertainty can be attributed to either lack of data or their inaccuracy, or an inexact mathematical model form. This inexact form can be due to either simplification or ignorance [6]. Model parameters cannot be estimated accurately when insufficient high-quality data are available. This uncertainty on the parameters can be reduced by collecting more data. Due to the asymptotic normality of the maximum likelihood estimator, a second-order description or confidence bounds of the statistical uncertainty is readily obtained by inverting the local Fisher information matrix [18].

Model inexactness can be introduced when a complex model is simplified and some of the variables are omitted from the analysis. Sometimes a nonlinear model is linearized or replaced by a response-surface model to save computer time. Uncertainties due to ignorance are by definition notoriously hard to quantify. They can only be described in probabilistic terms [12]. In this paper we will consider uncertainty due to missing variables and boundary conditions.

In comparison with Eq.(4.7), model uncertainty adds an extra layer of uncertainty to the problem, but does not fundamentally alter the structure of the problem. Let θ and M denote the model parameter and model form uncertainty, respectively. When the model uncertainties θ and M are not taken into consideration, the results obtained using

inherent variability x only can be considered conditional upon both the values used for the model parameters and the mathematical form used for the actual model itself: $y|\theta, M = g(x|\theta, M)$. The complete PDF model for the output $y = g(x, \theta, M)$ is:

$$
\begin{aligned}
f_{Y,\Theta,M}(y,\theta,M) &= f_{Y|\Theta,M}(y|\theta,M)\, f_{\Theta,M}(\theta,M) \\
&= f_{Y|\Theta,M}(y|\theta,M)\, f_{\Theta|M}(\theta|M)\, f_M(M)
\end{aligned} \tag{4.8}
$$

If the model uncertainties are independent of the inherent uncertainties the complete PDF model for the output $y = g(x, \theta, M)$ becomes separable and simplifies to:

$$
f_{Y,\Theta,M}(y,\theta,M) = f_Y(y) f_\Theta(\theta) f_M(M) \tag{4.9}
$$

5. Burgers's Equation. In this paper we consider a non-linear 1-D model problem, which includes both advection and diffusion. The viscosity is the only model parameter in this problem. The complete nonlinear Burgers equation

$$
\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = \frac{\partial}{\partial x}\left(\mu \frac{\partial u}{\partial x}\right) \tag{5.1}
$$

is a parabolic PDE which can serve as a model equation for the boundary-layer equations and is very similar to the equations governing fluid flow [3]. For simplicity the linearized equation

$$
\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = \frac{\partial}{\partial x}\left(\mu \frac{\partial u}{\partial x}\right) \tag{5.2}
$$

is often used. Equations (5.1) and (5.2) can be combined into a generalized equation (GBE)

$$
\frac{\partial u}{\partial t} + (c + bu)\frac{\partial u}{\partial x} = \mu \frac{\partial^2 u}{\partial x^2} \tag{5.3}
$$

For constant viscosity μ, $c = \frac{1}{2}$ and $b = -\frac{1}{2}$, the GBE (5.3) has the stationary solution

$$
u(x) = \tfrac{1}{2}\left[1 + \tanh\left(\tfrac{x}{4\mu}\right)\right] \tag{5.4}
$$

We solved the GBE as a steady-state boundary value problem with a second-order accurate central difference scheme using a conservative formulation for Eq.(5.3). The non-linear equation:

$$
\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}\left(\mu \frac{\partial u}{\partial x}\right) \tag{5.5}
$$

where the flux $f = \frac{u}{2}(1 - u)$, is solved iteratively. We drive the residual $r(u)$ to zero:

$$
r(u) = \frac{\partial f}{\partial x} - \frac{\partial}{\partial x}\left(\mu \frac{\partial u}{\partial x}\right) = 0 \tag{5.6}
$$

We implemented Newton's method on a uniform grid:

$$
\begin{aligned}
&\left(\tfrac{\partial r}{\partial u}\right)^{(n)} \Delta u^{(n)} = -r(u^{(n)}) \\
&u^{(n+1)} = u^{(n)} + \Delta u^{(n)}
\end{aligned} \tag{5.7}
$$

Using a second-order central difference scheme this leads to a tri-diagonal system of equations in each iteration:

$$
\begin{bmatrix}
b_1 & c_1 & & & & \\
a_2 & b_2 & c_2 & & 0 & \\
 & & \ddots & & & \\
 & 0 & & a_{n-1} & b_{n-1} & c_{n-1} \\
 & & & & a_n & b_n
\end{bmatrix}
\begin{Bmatrix}
\Delta u_1 \\ \Delta u_2 \\ \vdots \\ \Delta u_{n-1} \\ \Delta u_n
\end{Bmatrix}
= -
\begin{Bmatrix}
r(u_1) \\ r(u_2) \\ \vdots \\ r(u_{n-1}) \\ r(u_n)
\end{Bmatrix} \tag{5.8}
$$

where

$$\text{(5.9)} \qquad \begin{cases} a_j = \frac{\partial r_j}{\partial u_{j-1}} = -\frac{(1/2-u_{j-1})}{2\Delta x} - \frac{\mu_j}{\Delta x^2} \\ b_j = \frac{\partial r_j}{\partial u_j} = \frac{\mu_j+\mu_{j+1}}{\Delta x^2} \\ c_j = \frac{\partial r_j}{\partial u_{j+1}} = \frac{(1/2-u_{j+1})}{2\Delta x} - \frac{\mu_{j+1}}{\Delta x^2} \end{cases}$$

and $b_1 = b_n = 1$ and $a_1 = c_1 = a_n = c_n = r(u_1) = r(u_n) = 0$.

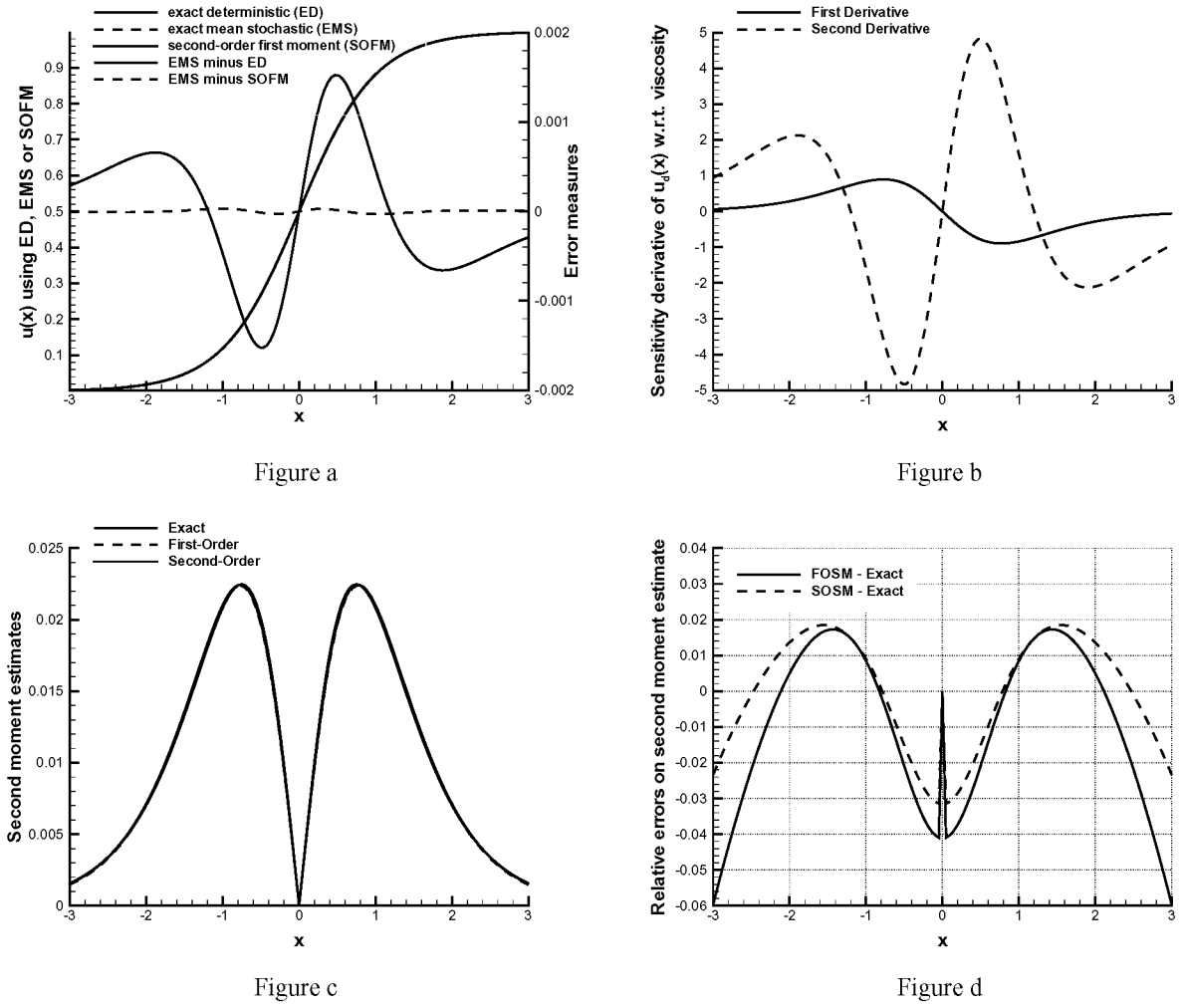


FIG. 6.1. *Comparison of exact and approximate analytical solutions to first and second moment of $u(x)$ using 129 grid points.*

6. Random Variable Modeling of the GBE. The differential equation Eq.(5.3) is solved for an assumed value of the viscosity μ, which does not depend on the independent variable x, i.e. μ is constant over the entire domain. The exact PDF of $u(x)$ can be obtained from the following equation [16]:

$$\text{(6.1)} \qquad f_U(u(x)) = \left| \frac{\partial}{\partial \mu} u(x,\mu) \right|^{-1} f_M(\mu)$$

from which the exact mean $\mathbb{E}[u(x)]$ and variance $\mathrm{Var}[u(x)] = \mathbb{E}[u^2(x)] - \mathbb{E}[u(x)]^2$ are readily computed using Mathematica [42]. These results are shown in Figure 6.1 for an assumed a Gaussian distribution for μ with mean value

$\mathbb{E}(\mu) = \overline{\mu} = 0.25$ and COV of 10%.

On the basis of the closed-form analytic solution in Eq.(5.4), we can compute analytic expressions for the first and second derivatives. We can readily compare first and second-order approximations with exact results. The first and second order approximation to the expected value $\mathbb{E}[u(x)]$ are given by:

$$\begin{aligned} \mathbb{E}_{FO}[u(x)] &= u(\overline{\mu}, x) \\ \mathbb{E}_{SO}[u(x)] &= u(\overline{\mu}, x) + \tfrac{1}{2}\mathrm{Var}(\mu)\, \tfrac{\partial^2 u(x)}{\partial \mu^2}\Big|_{\overline{\mu}} \end{aligned} \tag{6.2}$$

The first-order approximation is identical to the deterministic solution. The mean stochastic and second-order approximation thereof are shown on Figure 6.1a. The graphs almost coincide but the differences $\mathbb{E}[u(x)] - \mathbb{E}_{FO}[u(x)]$ and $\mathbb{E}[u(x)] - \mathbb{E}_{SO}[u(x)]$ are plotted along the Y-axis on the right side of Figure 6.1a. Since the error between the deterministic (a.k.a. first-order) and mean stochastic solution has a shape, which is very similar to the second-derivative (plotted in Figure 6.1b), it is to be expected that the second-order correction will lead to a much more accurate result. The error on the second-order first-moment (SOFM) estimate is 2 orders of magnitude smaller than the error on the first-order first-moment (FOFM) estimate. It is worth noting that it follows from Eq. (6.2) that, due to the symmetric nature of the Gaussian PDF, the derivatives of odd-order have no impact on the estimate for $\mathbb{E}[u(x)]$.

The first order approximation to the variance of $u(x)$ is:

$$\mathrm{Var}_{FO}[u(x)] \quad = \left(\tfrac{\partial u(x)}{\partial \mu}\Big|_{\overline{\mu}} \right)^2 \mathrm{Var}(\mu) \tag{6.3}$$

When μ is a Gaussian random variable, the second order approximation to the variance of u is:

$$\mathrm{Var}_{SO}[u(x)] \quad = \left(\tfrac{\partial u(x)}{\partial \mu}\Big|_{\overline{\mu}} \right)^2 \mathrm{Var}(\mu) + \tfrac{1}{2} \left(\tfrac{\partial^2 u(x)}{\partial \mu^2}\Big|_{\overline{\mu}} \mathrm{Var}(\mu) \right)^2 \tag{6.4}$$

Both the exact variance $\mathrm{Var}[u(x)]$ and the FO (6.3) and SO (6.4) approximations are even functions of x. Figure 6.1d compares the accuracy of the first-order and second-order solutions. It can be concluded that the second-order estimate, which is always greater than the first-order estimate for a Gaussian random variable, is generally speaking more accurate. Figure 6.2 shows the true PDF of $u(x)$ for selected values of x. The PDF shapes for $u(x_0)$ and $u(-x_0)$ are identical, but mirrored. Both the FOSM and SOSM estimate assume that the underlying distribution is symmetric. Consequently, the largest errors occur near the ends of the x-interval, i.e. where the PDF for $u(x)$ is strongly skewed. Since the PDF at $x = \pm 1$ is nearly Gaussian, the SOSM estimate is highly accurate. It is also important to note that a correct value $\mathrm{Var}[u(0)] \equiv 0$ is predicted by all methods. The skewness and kurtosis are shown in Figure 6.3. The kurtosis value $\kappa = 0$ at $x = 0$ is caused by the degenerate PDF for $u(0)$.

7. Gaussian Random Fields.

7.1. Second-Order Description. A continuous (or discrete) random field can be regarded as an infinite (or finite) set of random variables defined over its domain $\mathcal{X}$. One-dimensional random fields $Z(x)$ are also referred to as random processes. For all practical purposes, continuous random fields must be discretized for use in numerical algorithms, and we can concentrate on discrete random field characteristics only. Just like any other n-dimensional probability density, a random field is fully characterized either by its joint density or by its marginals and all its higher-order conditional PDFs, i.e of order $2, 3, \ldots, n-1$ [16]. In second-order analysis, a random field is described by the marginals and the auto-correlation function $R(x, x')$, which is usually defined as the correlation between the random variables $Z(x)$ and $Z(x')$ at two locations x and x' [40]:

$$R(x, x') = \frac{\mathrm{Covar}[Z(x), Z(x')]}{\sqrt{\mathrm{Var}[Z(x)] \cdot \mathrm{Var}[Z(x')]}} \tag{7.1}$$

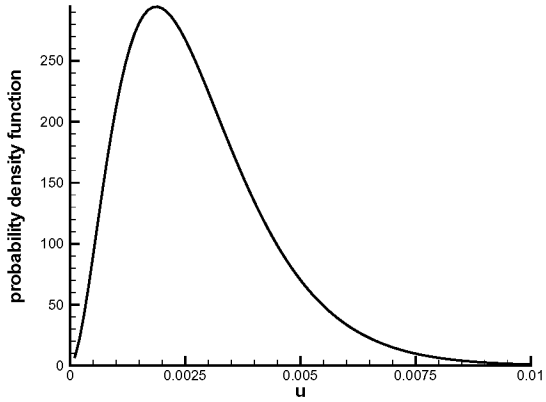


Figure a: $x = -3$

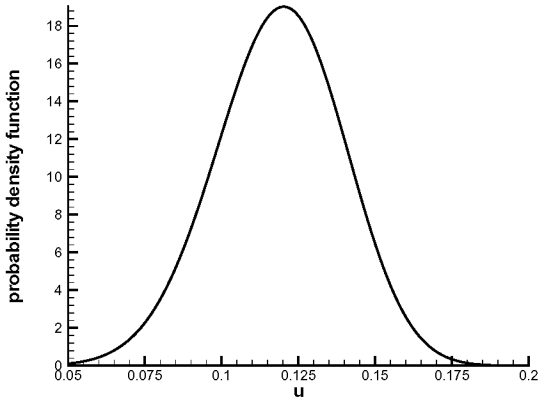


Figure b: $x = -1$

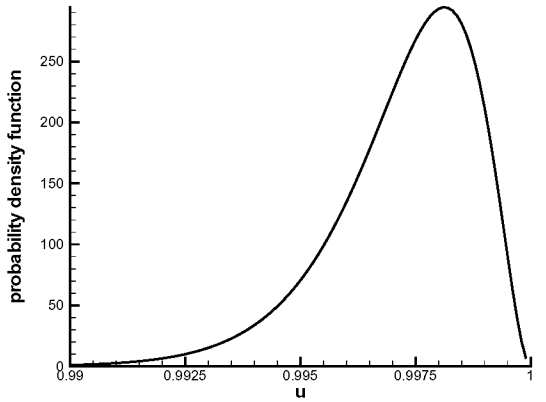


Figure c: $x = +3$

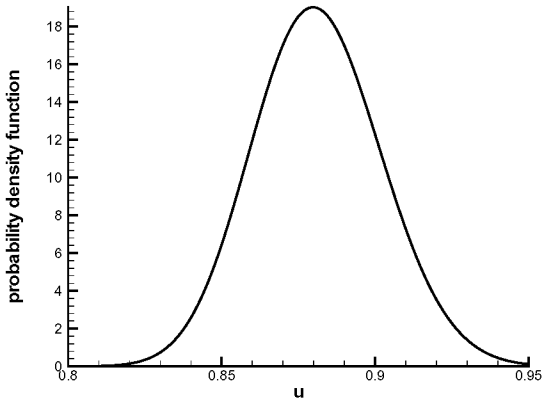


Figure d: $x = +1$

FIG. 6.2. *Exact PDF of $u(x)$ at selected values for x.*

For a homogeneous random field, the marginal distribution is independent of the location and the auto-correlation function R depends on only the distance or lag $\xi = x - x'$ between the two locations [29]. A homogeneous, Gaussian process is completely determined by its mean value $\mathbb{E}[Z(x)]$, variance $\mathrm{Var}[Z(x)]$ and auto-correlation $R(\xi)$.

TABLE 7.1
Commonly used families of auto-correlation models and corresponding spectral density functions [29].

Correlation model	Auto-correlation $R(\xi)$	Spectral density $S(\omega)$
exponential	$\exp(-\lvert\xi\rvert/l_c)$	$\frac{l_c}{\pi[1+(\omega l_c)^2]}$
squared exponential	$\exp\left[-(\xi/l_c)^2\right]$	$\frac{l_c}{2\sqrt{\pi}}\exp\left[-\left(\frac{\omega l_c}{2}\right)^2\right]$
triangular	$\begin{cases} 1-\frac{\lvert\xi\rvert}{l_c} & \text{if } \lvert\xi\rvert \le l_c \\ 0 & \text{if } \lvert\xi\rvert > l_c \end{cases}$	$\frac{1-\cos(\omega l_c)}{\pi\omega^2 l_c}$
damped sinusoidal	$\frac{\sin(\pi\lvert\xi\rvert/l_c)}{\pi\lvert\xi\rvert/l_c}$	$\begin{cases} \frac{l_c}{2\pi} & \text{if } \lvert\omega\rvert \le \frac{\pi}{l_c} \\ 0 & \text{if } \lvert\omega\rvert > \frac{\pi}{l_c} \end{cases}$

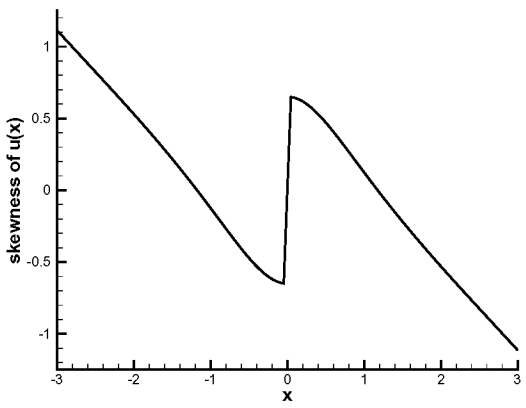


Figure a: compare with $\gamma_1 = 0$ for Gaussian PDF

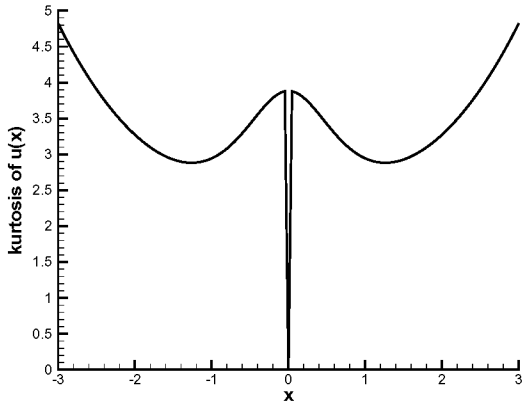


Figure b: compare with $\kappa = 3$ for Gaussian PDF

FIG. 6.3. *Skewness and kurtosis of the exact PDF of $u(x)$ using 129 grid points.*

Some commonly used models for the auto-correlation function are listed in Table 7.1 and shown in Figure 8.2a. The first two are perhaps most commonly encountered in the literature [21]. The first one applies to a non-differentiable field, whereas the second one is found in differentiable fields. The correlation length l_c is a very important measure for the rate of fluctuation within the random field model. A larger correlation length l_c corresponds to a slower varying random field as is illustrated in Figure 7.1 for the exponential correlation model.

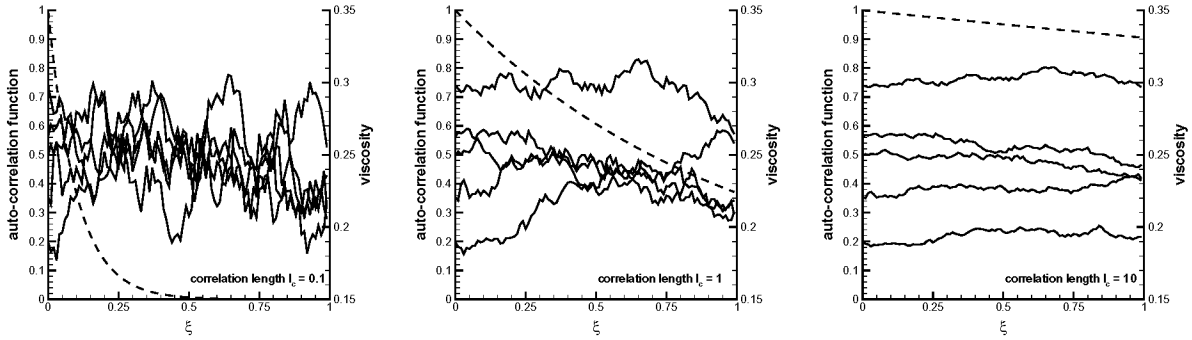


FIG. 7.1. *Exponential auto-correlation function and sample random fields for different correlation length l_c. Marginal density for the viscosity is Gaussian with mean value 0.25 and COV= 10%.*

7.2. Discretization of Random Fields. Most modern numerical solution algorithms to PDE's (such as finite element, finite volume or finite difference) require that all continuous parameter fields are discretized. When spatial or temporal uncertainty effects are directly included in the analysis it makes sense to use random fields for a more accurate representation of the variations. For instance, a thickness of the wing-skin is random but is also not likely to be exactly the same everywhere. The spatial or temporal fluctuations of a parameter cannot be accounted for if that parameter is modeled by only a single random variable.

We only give a brief overview of some of the discretization methods, a more extensive review is available in the literature [34]. Some of these methods are limited to Gaussian fields only. We will discuss:

1. midpoint method
2. locally averaged method
3. shape function method
4. series expansion method

Other methods include locally averaged subdivision [10] and the turning-bands method [9].

The simplest method of discretization is the midpoint method [7]. In this method the field within the one element or cell is described by a single random variable, which represents the field value at the centroid of the element. The value of the field is assumed to be constant within the entire cell. Consequently, the discretized field is step-wise constant with discontinuities at the cell boundaries. Its statistics (mean, variance, ...) are readily computed from the statistics at the centroids.

In the locally averaged method, the field within each cell is described in terms of the spatial average of the field over the element [40]. The discretized field is still constant over each cell with step-wise discontinuities at the cell boundaries, but one can expect a better fit due to the averaging process. Typically, the variance of the averaged random process is smaller than the local variance of the random field, used for midpoint discretization. Expressions for the statistics will be derived in Section 9.

The shape function method [23] describes the random field using shape functions anchored at a set of nodal values. The method is particularly elegant and efficient in conjunction with the finite element method [22]. The random field description is continuous across cell boundaries and this represents a clear improvement over the midpoint and locally averaged discretization methods.

Yet another method uses a series expansion approach to discretize the random field. Legendre basis functions are used with Gaussian random fields in [20]. Other functions can be used in conjunction with different marginal PDFs [43]. Provided that the exact eigenfunctions in the Karhunen-Loève expansion are available, this method requires the smallest number of random variables for a given level of accuracy [39]. However, often there is no exact solution to the integral eigenvalue problem associated with the Karhunen-Loève expansion. In this case, the resulting approximations essentially reduce the series expansion method to a shape function method [21].

In this report, we compare results using the midpoint and locally averaged random field discretization methods for Monte Carlo simulations. Polynomial chaos results will be reported in an upcoming report [41].

8. Generation of Midpoint Random Field Samples. Generally speaking, three families of methods can be used to generate sample random fields. The different methods will be applied only in conjunction with the auto-correlation functions given in Table 7.1, but are – in most cases – applicable to other auto-correlation functions as well. The random field can be discretized over the grid, which is used to solve the differential equation, or over any other grid. Typically, the random field grid is a subset of the finite element/difference/volume grid but this – strictly speaking – need not be the case. Only Gaussian random fields are included in this discussion.

8.1. Auto-regressive processes. Processes with an exponentially decaying auto-correlation function can easily be transformed into first-order auto-regressive, or AR(1) processes, which are also known as Markov processes. In such processes the auto-correlation decreases steadily with increasing distance between sample points. For a uniform grid the discrete AR(1) process $Z(x)$ is defined by:

$$Z(x_i) = \phi Z(x_{i-1}) + a_i \tag{8.1}$$

where $|\phi| < 1$ and all a_i are independent, identically distributed zero-mean Gaussian variables. The discrete process has the following second-order characteristics [35]:

$$\begin{aligned} &\mathbb{E}[Z(x)] = 0 \\ &\mathrm{Var}[Z(x)] = \tfrac{\mathrm{Var}[a]}{1-\phi^2} \\ &R_Z(x_i, x_{i\pm k}) = \phi^{|k|} \end{aligned} \tag{8.2}$$

If we chose the parameters

$$\begin{aligned} &\phi = \exp(-1/l_c) \\ &\mathrm{Var}[a_i] = 1 - \exp(-2/l_c) \end{aligned} \tag{8.3}$$

it readily follows that the field $Z(x)$ has an exponential correlation function $R_Z(x_i, x_{i\pm k}) = R_Z(\xi) = \exp(-|\xi|/l_c)$, where $\xi = x_i - x_{i\pm k}$.

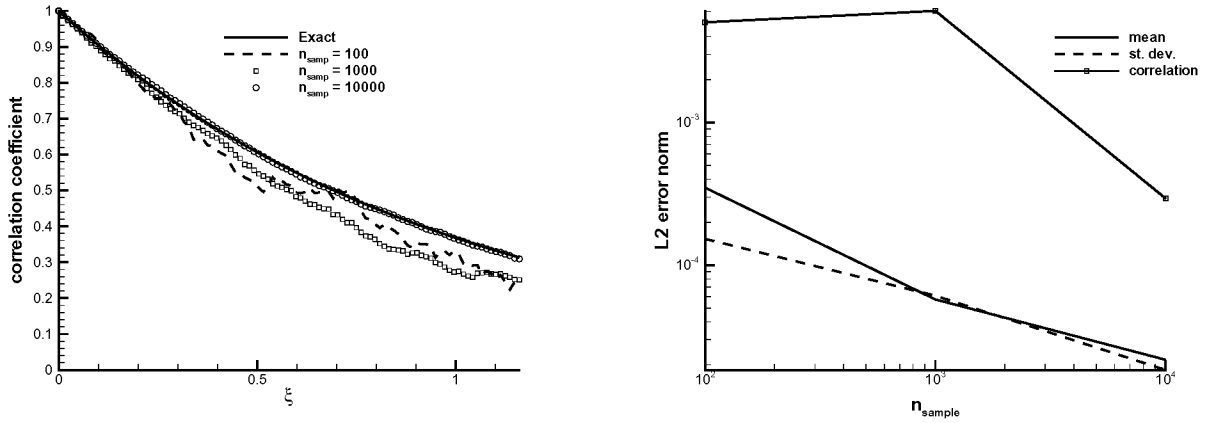


Figure a: correlation function

Figure b: L2 Norm of error estimate

FIG. 8.1. *Theoretical and sample auto-correlation functions for exponential correlation model with* $l_c = 1$.

TABLE 8.1

Mean Relative (in percent) and Absolute Squared Error on the sample statistics for the exponential correlation model with $l_c = 1$. *Marginal density for the viscosity is Gaussian with mean value 0.25 and COV*$= 10\%$.

	Relative error (in %) on			Absolute error (L2 norm) on		
$n_{samples}$	Mean	St. Dev.	Correl.	Mean	St. Dev.	Correl.
100	0.14	0.61	1.13	0.00035	0.00015	0.00503
1000	0.02	0.24	1.43	0.00006	0.00006	0.00605
10000	0.01	0.07	0.06	0.00002	0.00002	0.00029

For uniform grids, the sample auto-regressive processes are easily generated using this method. The convergence of the sample statistics to the true values of the random field is shown in Figure 8.1 and Table 8.1. From a practical point of view, the stability and accuracy of the ensemble statistics is very important for the simulation process. A more general description of auto-regressive processes can be found in [33] and [27].

8.2. Covariance decomposition. In this method, the covariance matrix of the discretized field is generated first. For a zero-mean, homogeneous process $Z(x)$ the symmetric, positive definite covariance matrix is directly derived from the auto-correlation function:

$$\text{Covar}[Z(x_i), Z(x_j)] = R(\xi)\text{Var}(Z) \tag{8.4}$$

where $\xi = |x_i - x_j|$ is the distance between the points x_i and x_j.

Along with the mean value $\mathbb{E}[Z(x)] = \overline{Z}$, the covariance matrix fully describes the discretized Gaussian random field. Using the covariance decomposition method the sample values $Z_j(x_i)$ of $Z(x_i)$, $i = 1 \ldots n$ can be generated on the basis of n independent random normal variables y_i:

$$\{Z(x_i)\} = \{\overline{Z}\} + [\mathbf{A}]^T . \{y_i\} \tag{8.5}$$

where the matrix $[\mathbf{A}]$ contains the eigenvectors of the covariance matrix. The variance of each of the independent Gaussian variables y_i is equal to λ_i, the i^{th} eigenvalue of the covariance matrix. In this method the sample mean and covariances will only approximately match the required values and approach the true values as the number of samples increases.

It is possible to ensure that each ensemble matches the prescribed means and covariances, even with a small number of samples. This preconditioning method [44] combines the modal decomposition of the covariance matrix with the spectral representation of the random field, which is described in the next section. However, in this method the variates y_i or $Z(x_i)$ are not directly generated from a Gaussian distribution; they are only approximately Gaussian by means of the central limit theorem. They are generated as a sum of independent and identically distributed random variables, which approaches a Gaussian random variable if sufficient terms are included in the sum.

The covariance decomposition method readily identifies the most important random variables as the ones with the largest eigenvalues. This is useful if one wants to reduce the number of random variables in the random field description.

8.3. Spectral representation. Unlike the previous methods, the representation of a random field using the spectral method is continuous across the elements. The underlying idea is that a continuous random function can be expanded on a basis of deterministic functions with random coefficients. The goal is to achieve a sufficiently accurate representation of the random field with a finite, i.e. discrete, number of functions and coefficients.

The spectral density $S(\omega)$ (or power spectrum) of a stationary process $Z(x)$ is given by the Fourier transform of its auto-correlation $R(\xi)$:

$$S(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} R(\xi) \exp(-j\omega\xi) d\xi \tag{8.6}$$

Since $R(\xi)$ is an even function for a real-valued process $Z(x)$, it follows that the $S(\omega)$ is a real function of ω:

$$S(\omega) = \frac{1}{\pi} \int_{0}^{\infty} R(\xi) \cos(\omega\xi) d\xi \tag{8.7}$$

The random field discretization is achieved through the discretization of the spectral density. The power spectrum is approximated at some pre-selected discrete frequencies, typically chosen equally spaced in $[-\omega_{\max}, \omega_{\max}]$.

$$Z_j(x) = \overline{Z} + \sum_{i=0}^{N_{freq}} \sqrt{2\text{Var}[Z(x)]S(\omega_i)\Delta\omega} \cos(\omega_i' x + \phi_i) \tag{8.8}$$

where $\Delta\omega$ is the spacing between the sampling frequencies, ω_i' is a small perturbation of ω_i and ϕ_i is the phase angle [37].

The process generated according to Eq.(8.8) is asymptotically Gaussian as $N_{freq} \to \infty$. Since the highest frequencies are ignored in the discretization of the random field, part of the variance of the process will be missed. Another source of discretization error stems from the finite number of terms used in the representation. The convergence to the target auto-correlation $R(\xi)$ or to the target spectral density $S(\omega)$ is inversely proportional to N_{freq}^2 [38]. Additional details of this method are described in [37].

If the eigenvalues and eigenfunctions of the covariance matrix Eq.(8.4) are known, the Karhunen-Loève expansion requires the smallest number of functions to achieve a prescribed level of accuracy. However, in many cases the eigenfunctions are only approximately known and the approximate K-L expansion is no more efficient than the shape function expansion [22]. Other expansions may require more terms to achieve the same accuracy but require substantially less computational effort. A detailed discussion of the truncation error in each of the methods can be found in [21] and [45].

The slow fluctuations are represented using the lowest frequencies, whereas the high frequencies are required to accurately model the rapid fluctuations. It automatically follows that much higher frequencies are required in Eq.(8.8) to accurately model the sample random field in Figure 7.1a than for the one in Figure 7.1c. Figure 8.2 illustrates the relationship between the auto-correlation model and the frequency range which is required.

9. Locally Averaged Random Fields. For the 1-D problem considered here we get the following statistics for the locally averaged random field:

$$Z_X(x) = \frac{1}{X} \int_{x-X/2}^{x+X/2} Z(s) ds \tag{9.1}$$

where X is the cell length and x is the midpoint of the cell. Since the field is homogeneous,

$$\mathbb{E}(Z_X(x)) = \mathbb{E}(Z(x)) = \overline{Z} \tag{9.2}$$

The averaging process typically reduces the variance [40]. The variance function $\gamma(X)$ measures the reduction of the point variance Var$[Z]$ due to the averaging process. For a homogeneous random field::

$$\text{Var}[Z_X] \equiv \sigma_X^2 = \gamma(X)\text{Var}[Z] \tag{9.3}$$

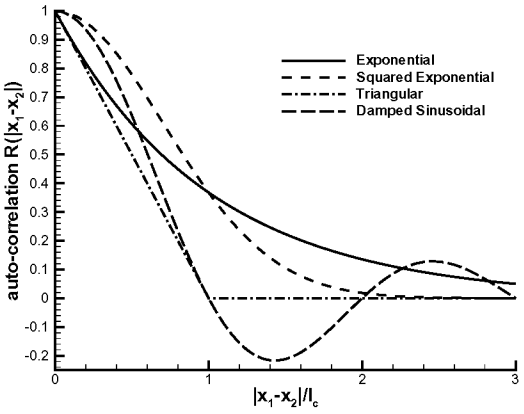


Figure a: Auto-correlation models

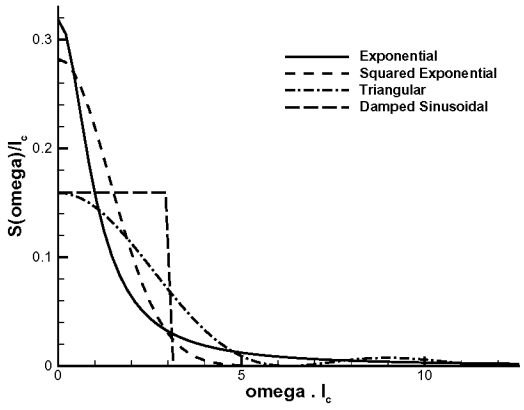


Figure b: Power spectral density

FIG. 8.2. *Auto-correlation functions and corresponding spectral densities used in simulations.*

and, for a homogeneous field, the variance function solely depends on the cell size X:

$$\gamma(X) = \frac{2}{X}\int_0^X \left(1 - \frac{\xi}{X}\right) R(\xi)d\xi \tag{9.4}$$

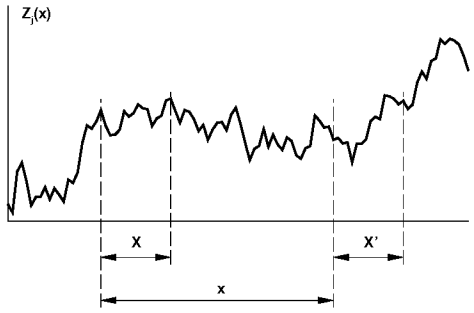


FIG. 9.1. *Definition of intervals and distances used in Eq.(9.7).*

The scale of fluctuation θ is defined as follows:

$$\theta = \lim_{X\to\infty} X\gamma(X) \tag{9.5}$$

which indicates that, asymptotically for large averaging length X: $\gamma(X) \sim \theta/X$. Combination of Eqs.(9.5) and (9.4) implies that

$$\theta = 2\int_0^\infty R(\xi)d\xi = \int_{-\infty}^\infty R(|\xi|)d\xi \tag{9.6}$$

It follows immediately from the spectral density definition in Eq. (8.6) that the scale of fluctuation is proportional to the ordinate $S(\omega = 0)$ of the spectral density function $\theta = 2\pi S(0)$.

The scale of fluctuation provides a lot of information over the extent of the correlation within the random field. If the correlation dies out rapidly, the asymptotic relationship will be achieved for a relatively small averaging length X. If two locally averaged field values are more than distance θ apart, Eq.(9.5) indicates that they are almost independent of each other.

On the basis of Eq. (9.4), the second-moment properties of the locally averaged random field can be computed. These can then be used to generate sample random fields using the covariance decomposition method. The correlation coefficient $\rho_{Z_X,Z_{X'}}$ between the locally averaged random field values Z_X and $Z_{X'}$ is given by:

$$\rho_{Z_X,Z_{X'}} = \frac{\Delta(x-X) + \Delta(x+X) - 2\Delta(x)}{2\Delta(X)} \tag{9.7}$$

where $\Delta(\bullet) = (\bullet)^2\gamma(\bullet)$ and the distances x and $X = X'$ are defined in Figure 9.1.

10. Random Field Modeling of the GBE. In reality, the viscosity also depends – among other factors – on the temperature, for which no information is available in the Burgers model (5.3). The model uncertainty caused by these missing variables is best described using a random field, which can be thought of as an infinite set of random variables,

each one of them describing the variability of the viscosity at a particular location x. The second-order description of a homogeneous random field [29] is given by the marginal PDF $f(\mu)$ and the auto-correlation function $R(\xi)$ which are both independent of the location x for a homogeneous random field:

$$R(\xi) = \frac{\text{Covar}[\mu(x), \mu(x+\xi)]}{\text{Var}[\mu(x)]} \tag{10.1}$$

where $\xi = x_2 - x_1$ is the lag or distance between the two locations considered [40]. It is assumed that the marginal density $f(\mu)$ is Gaussian with mean value $\overline{\mu} = 0.25$ and a COV of 10%.

The impact of the uncertain boundary conditions on the solution is discussed in the next section. This uncertainty is ignored in this section: we directly applied the Dirichlet boundary conditions, obtained from the deterministic steady-state solution Eq.(5.4):

$$u(\pm x_{\max}) = \tfrac{1}{2}\left[1 + \tanh\left(\tfrac{\pm x_{\max}}{4\mathbb{E}(\mu)}\right)\right] \tag{10.2}$$

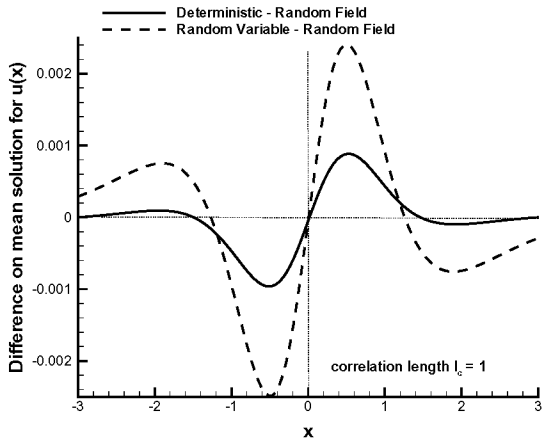


FIG. 10.1. *Difference between deterministic, mean random variable and mean random field solution (exponential auto-correlation with $l_c = 1$).*

The results in Section 6 using a random variable model reveal that the difference between the deterministic and the mean stochastic solution is small from a practical point of view. This is still the case for the random field model as well. Figure 10.1 indicates, however, that the difference between the deterministic and the mean random field solution is somewhat smaller than the difference between the mean random variable and the mean random field solution. To a large extent, this can be attributed to the formulation of the boundary conditions using Eq.(10.2); which are identical for the deterministic and random field solution. Note that the mean random variable solution is an exact result, whereas the mean random field solution is the results of 10^6 Monte Carlo simulations. The standard error on the Monte Carlo results is less than 0.00006.

For this reason we will not study the impact of the correlation length l_c and auto-correlation model $R(\xi)$ on the mean solution. The discussion is limited to the effects of the random field modeling on the standard deviation only. It immediately follows from Figure 7.1 that the random field and random variable models lead to identical solutions (shown in Figure 6.1c) for an infinite correlation length $l_c = \infty$.

10.1. Comparison of midpoint and locally averaged solutions. When the viscosity $\mu(x)$ is represented by a random field, the numerical solution of the GBE in Eq.(5.3) requires a discretization at two different levels:

1. discretization required to solve the differential equation and compute an accurate solution for the velocity field $u(x)$.

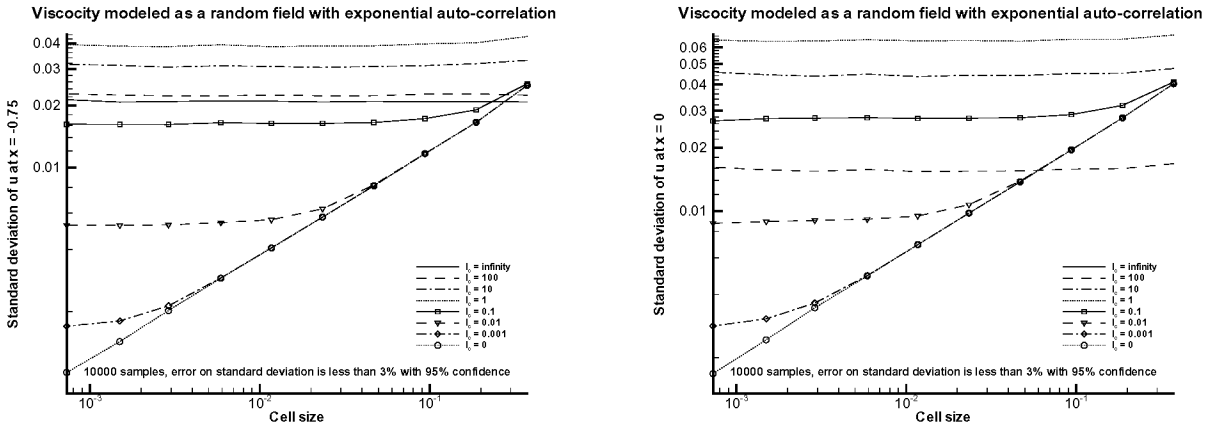

Viscocity modeled as a random field with exponential auto-correlation
Standard deviation of u at x = -0.75
0.04
0.03
0.02
0.01
10000 samples, error on standard deviation is less than 3% with 95% confidence
Cell size
Viscocity modeled as a random field with exponential auto-correlation
Standard deviation of u at x = 0
0.06
0.05
0.04
0.03
0.02
0.01
10000 samples, error on standard deviation is less than 3% with 95% confidence
Cell size

Figure a:$x = -0.75$ Figure b: $x = 0$

FIG. 10.2. *Cell size dependence of the variance using the midpoint discretization method.*

2. discretization required to model the fluctuations within the random field $\mu(x)$ and ensure that the resulting discretized equations are an accurate representation of the stochastic differential equation.

A more detailed error analysis of the second-order central difference scheme used in the analysis is documented elsewhere [41]. In this paper, the focus is on the discussion of the required discretization level of the random field itself. All the results in this Section apply to the exponential correlation model, but the same conclusions can be drawn for other correlation models.

In this section we compare Monte Carlo results obtained using either the midpoint or the locally averaged discretization method. The sample random fields for the midpoint method are generated using the auto-regressive property of the exponential correlation model Eq.(8.1). The sample random fields for the locally averaged method are obtained using covariance decomposition Eq.(8.5).

It is intuitively clear [7] that the midpoint discretization method tends to overestimate the variability of the output quantities of interest if the discretization is too coarse. This can be explained as follows. When the cell size is much larger than the correlation length l_c or the scale of fluctuation θ, the viscosity fluctuates rapidly with the cell (like in Figure 7.1a). Consequently, the midpoint values of the different cells will be practically independent of each other. As a result thereof, Var$[u(x)]$ will decrease if the number of cells increases. For uncorrelated midpoint values Var$[u(x)]$ will be inversely proportional to the number of cells. For a non-zero correlation length l_c, the fluctuations virtually disappear when the cell size is smaller than the correlation length l_c (like in Figure 7.1c). In this case the midpoint values of neighboring cells become highly correlated. Dividing the cells even further leads to basically identical midpoint values for $\mu(x)$ in the newly created cells and Var$[u(x)]$ stabilizes.

Figure 10.2 illustrates this effect for various correlation lengths l_c and an exponential correlation model for two locations:

- $x = -0.75$: this is near the location where the random variable model (Figure 6.1c) has its largest variance.
- $x = 0$: this is the location where the random variable model (Figure 6.1c) has zero variance.

When μ is modeled as a random variable, the Burgers equation is solved once for an assumed value of the viscosity μ. The mathematical expression of the closed form solution for $u(x)$, given in Eq.(5.4), is independent of the actual value of μ. When the viscosity is modeled as a random field, however, the ODE in Eq.(5.3) becomes a stochastic differential equation. It is immediately clear from Figure 10.2b that the solution of the stochastic differential equation

has entirely different properties than Eq.(5.4). The random variable solution of Eq.(5.4) results in $\text{Var}[u(0)] \equiv 0$. Figure 10.2b indicates that this result is recovered in the limit for $l_c = \infty$, but that this is generally not the case for a finite correlation length.

Figure 10.2 shows that the standard deviations stabilizes once the cell size is roughly equal to the scale of fluctuation θ. Recall that $\theta = 2l_c$ for the exponential auto-correlation model. It is interesting to note that the white noise random field model ($l_c = 0$) converges to a deterministic response, i.e. $\text{Var}[u(x)] \equiv 0$.

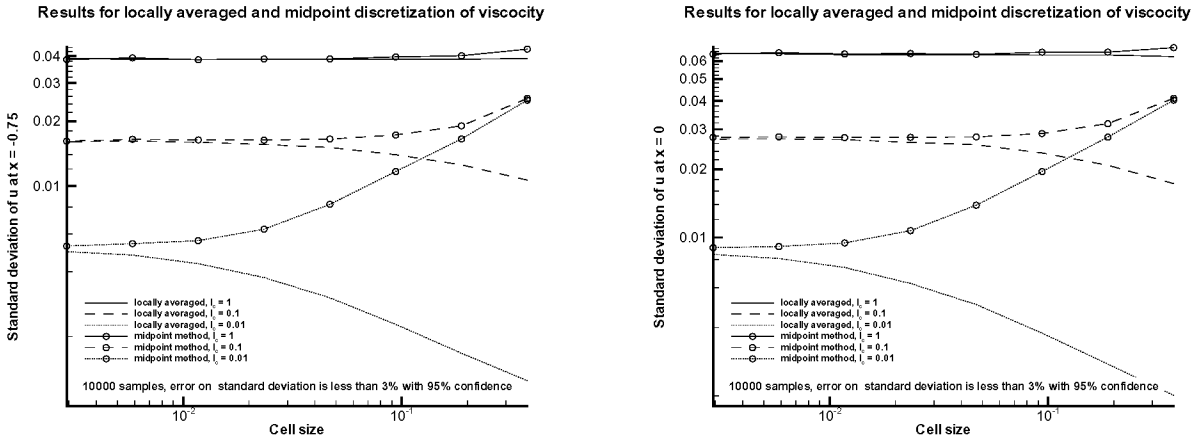

Results for locally averaged and midpoint discretization of viscocity
Standard deviation of u at x = -0.75
0.04
0.03
0.02
0.01
locally averaged, l_c = 1
locally averaged, l_c = 0.1
locally averaged, l_c = 0.01
midpoint method, l_c = 1
midpoint method, l_c = 0.1
midpoint method, l_c = 0.01
10000 samples, error on standard deviation is less than 3% with 95% confidence
10^-2
10^-1
Cell size
Results for locally averaged and midpoint discretization of viscocity
Standard deviation of u at x = 0
0.06
0.05
0.04
0.03
0.02
0.01
locally averaged, l_c = 1
locally averaged, l_c = 0.1
locally averaged, l_c = 0.01
midpoint method, l_c = 1
midpoint method, l_c = 0.1
midpoint method, l_c = 0.01
10000 samples, error on standard deviation is less than 3% with 95% confidence
10^-2
10^-1
Cell size

FIG. 10.3. *Comparison of cell size dependence of the variance using the midpoint and the locally averaged discretization method.*

On the other hand, locally averaged discretization causes the true variance to be underestimated [40]. When the cell size is much larger than l_c (as in Figure 7.1a) the local averaging process drastically reduces the variance as explained in Section 9 (variance function γ close to 0). Consequently, the variance of the discretized viscosity field in Eq.(5.3) will be underestimated. When the cell size is much smaller than l_c (as in Figure 7.1a) the local averaging process leaves the variance virtually intact (variance function γ close to 1) and the full variance of the process is included in the discretized random field. Figure 10.3 indicates that a stable solution for the variance is obtained if the cell size is smaller than the correlation length l_c and is approached from below as the cell size decreases. Figure 10.3 also shows that the midpoint and locally averaged results will tend to bracket the true variability of the output quantity of interest.

Figure 10.4 summarizes the impact of the correlation length l_c on $\text{Var}[u(x)]$. It follows from the figure that for practical purposes it suffices to only know the order of magnitude of l_c. For $l_c \to \infty$, the random field is represented by a single random variable and we obtain the same solution as in Figure 6.1c.

10.2. Effect of auto-correlation function or spectral density. In this section, the effect of the auto-correlation model on the standard deviation is assessed. All Monte Carlo solutions are computed using midpoint discretizations of the viscosity field. The grids are fine enough to obtain converged values for the statistics. All random field samples are generated using the spectral representation method. The power spectral densities were truncated at $\omega_{\max}$, corresponding to the 95%-level:

$$2\int_0^{\omega_{max}} S(\omega)d\omega = 0.95 \tag{10.3}$$

For all computations the frequency range $[-\omega_{\max}, \omega_{\max}]$ is discretized using 100 evenly spaced frequencies; the truncation frequency for each model is listed in Table 10.1. The standard deviations computed for two different assumed correlation lengths $l_c = 0.1$ and $l_c = 10$ using all 4 auto-correlation models are shown in Figure 10.5.

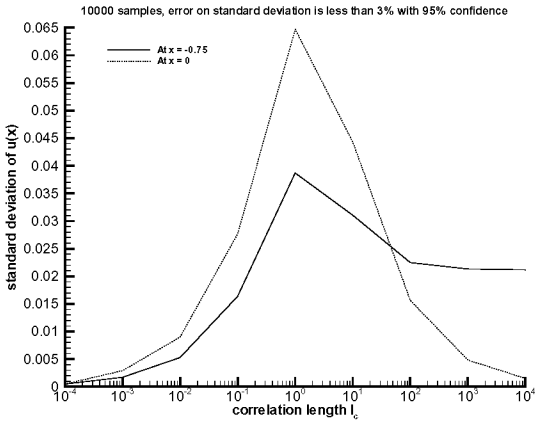

10000 samples, error on standard deviation is less than 3% with 95% confidence
At x = -0.75
At x = 0
standard deviation of u(x)
correlation length l_c

FIG. 10.4. *Impact of the correlation length l_c on the standard deviation (midpoint discretization method, all results are grid-converged)*

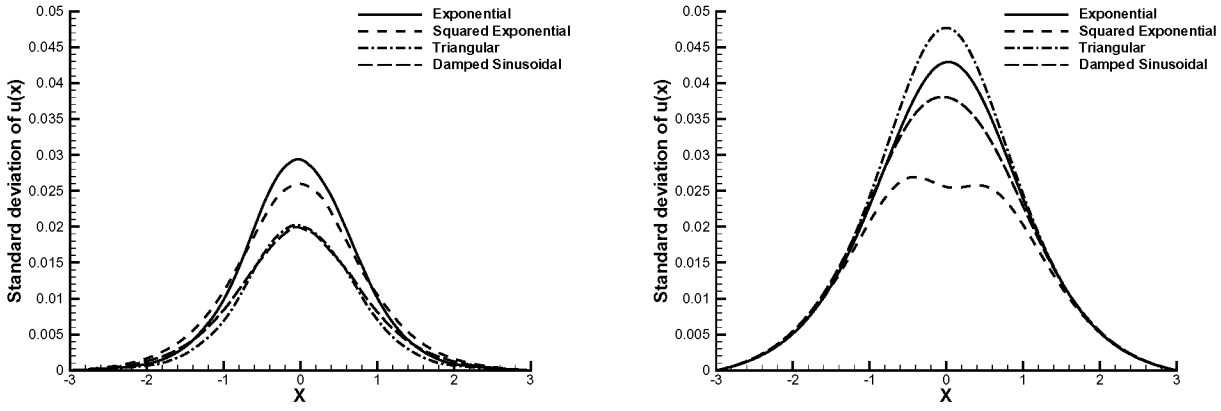


Figure a: $l_c = 0.1$ Figure b: $l_c = 10$

FIG. 10.5. StDev$[u(x)]$ *obtained using 1000 midpoint Monte Carlo simulations of the random field and 512 grid cells.*

As suggested by Figure 10.4 for the exponential auto-correlation model, Figure 10.5 shows that for all correlation kernels used the variability is generally larger for $l_c = 10$ than for $l_c = 0.1$. For the shorter correlation length $l_c = 0.1$ the triangular and damped sinusoidal correlation model lead to nearly identical variability in $u(x)$ (Figure 10.5a). Figure 10.5b shows that this is by no means a universal trend, but depends on the correlation length l_c. It can be

TABLE 10.1
Truncation frequency $\omega_{\max}$ used for computation of results in Figure 10.5.

Correlation model	$\omega_{\max} l_c$
exponential	12.7
squared exponential	3.92
triangular	13.0
damped sinusoidal	2.98

concluded that for $l_c = 0.1$, the fluctuations in the viscosity field are too rapid for the weaker medium-to-long-range correlations, caused by oscillating behavior of the correlation function (see Figure 8.2a), to have an effect. The weak correlations between points separated by a distance larger than l_c are drowning within the rapidly fluctuating field. When the viscosity field is varying more slowly, these correlations are strong enough to have an impact on the solution. Note that the slight asymmetry in Figure 10.5b (most pronounced for the squared exponential model) is entirely statistical in nature. Only 1000 Monte Carlo simulations were used to compute the results; the asymmetry disappears when more simulations are performed.

11. Uncertain Boundary Conditions. In the previous section, deterministic Dirichlet boundary conditions were imposed. In many CFD applications the differential equations have to be integrated over a (semi)-infinite domain. In practice, the integration domain is cut off at some "far field" boundary and deterministic boundary conditions are imposed. This approach essentially ignores all variability outside the truncated integration domain. In this section, the impact of the uncertainty modeling of the boundary conditions on the second-order statistics of the solution in the interior of the integration domain is analyzed.

In many practical cases, an exact, statistical treatment of the effects of the truncation of the integration domain will be impossible. Also, the data required to build the sophisticated statistical model for such an analysis may be unavailable. Therefore, we focus on a more practical second-order treatment of these uncertainties.

In particular we will compare the following 4 models:

(a) apply deterministic boundary conditions obtained from the steady state solution

(b) apply stochastic boundary conditions obtained from the closed-form steady state solution

(c) apply approximate stochastic boundary conditions, ignoring correlation effects

(d) apply approximate stochastic boundary conditions, accounting for the correlation between the boundaries

Model (a) completely ignores boundary condition uncertainty. Since a closed-form solution exists for this particular problem if the viscosity is modeled by a single random variable, we can get a fairly accurate modeling of the boundary condition uncertainty. For Model (b), we applied the exact PDF obtained for the boundary conditions $u(-x_{\max})$ and $u(x_{\max})$ to the random field model. This Model (b) still represents an approximation in the sense that the random field uncertainties are all lumped together on the boundary.

In many practical cases we may only have second moment information about the boundary condition uncertainty (i.e., mean value and variance). It then makes sense to add a random perturbation to the deterministic boundary conditions at $\pm x_{max}$. This is done in Models (c) and (d). Since $0 \leq u(x) \leq 1$, a Gaussian perturbation at the boundary can physically not be justified. Since only very scant information may be available regarding the uncertainty at the boundary conditions, we used the following basic perturbation model, based on the deterministic boundary conditions Eq.(10.2):

$$\begin{cases} u(-x_{\max}) = 0 + \delta_1 \\ u(x_{\max}) = 1 - \delta_2 \end{cases} \tag{11.1}$$

where δ_i are exponentially distributed random variables, with variance equal to the variance of the PDF in Model (b). Model (c) ignores the correlation effects between δ_1 and δ_2; while these are taken into consideration in Model (d). The correlated exponentially distributed random variables are generated using the Nataf transformation [8]. Figure 11.2 compares the approximate exponential marginal PDF for $u(-x_{\max})$ used in Models (c) and (d) with the PDF used in Model (b).

Figure 11.3 shows that, when the uncertainty on the boundary conditions is included, the difference between the deterministic and the mean stochastic solution is three times larger than when the uncertainty on the boundary conditions is not included (see Figure 10.1). For practical purposes, this difference remains small though ($l_c = 1$).

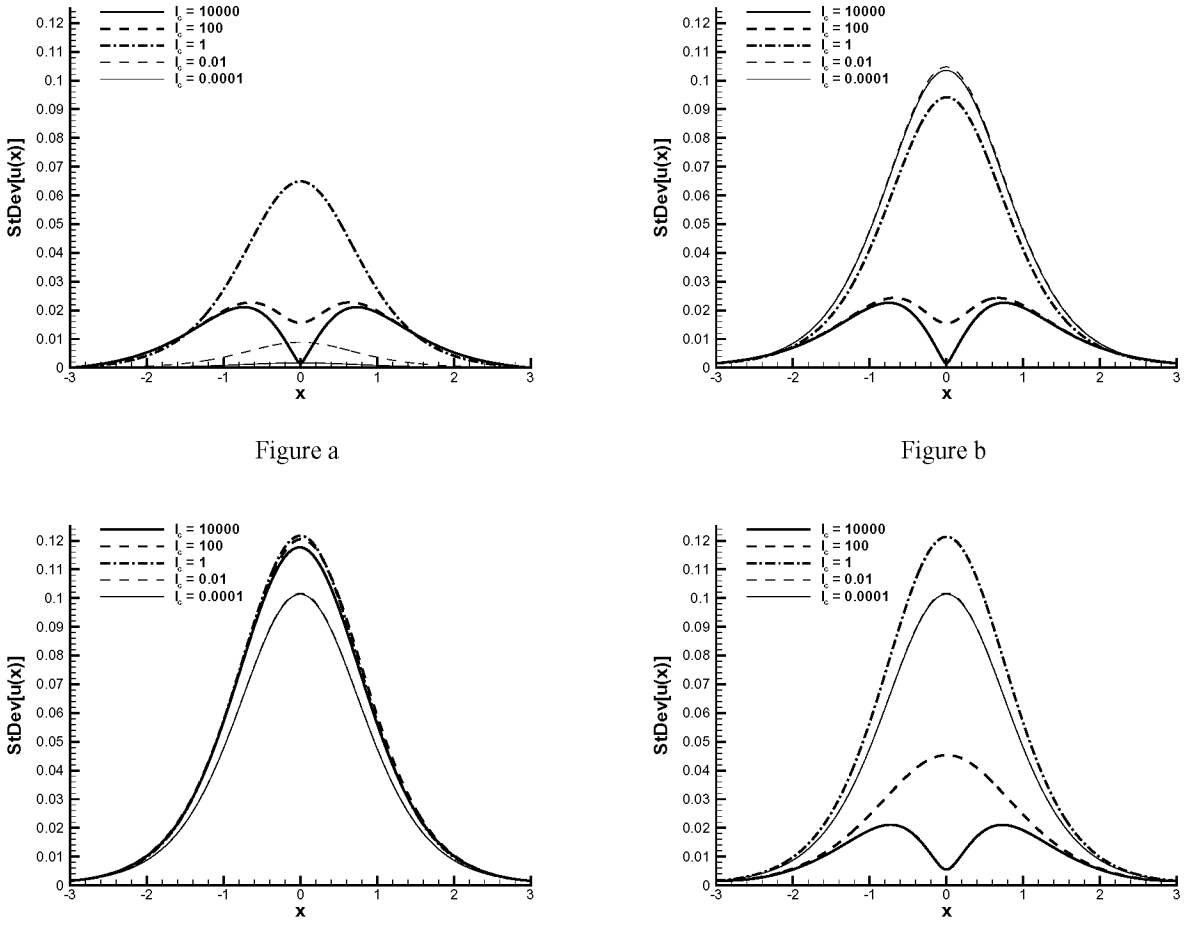


Figure a

Figure b

Figure c

Figure d

FIG. 11.1. *Comparison of* StDev$[u(x)]$ *using different boundary condition models and 8193 grid points.*

Note that the mean random variable solution is an exact result, whereas the mean random field solution is the results of 10^6 Monte Carlo simulations. The standard error on the Monte Carlo results is less than 0.00006.

Regarding the standard deviation of $u(x)$, the following conclusions can be drawn from the analysis (see Figures 11.1 and 11.4):

- Ignoring the uncertainty of the boundary conditions grossly underestimates the variability in the interior domain when the correlation length l_c is substantially shorter than the integration domain. Note that this will usually be the case; otherwise there is no reason to use a random field model.
- The importance of an accurate estimation of the correlation between random variables is clear from a comparison of Models (c) and (d). If the correlation length is long and substantial correlation is present between δ_1 and δ_2, Models (c) and (d) give drastically different results.
- Even though the exponential PDF is a crude approximation of the boundary uncertainty (see Figure 11.2), the variance of Model (d) follows the trends of Model (b) remarkably well for all l_c-values.

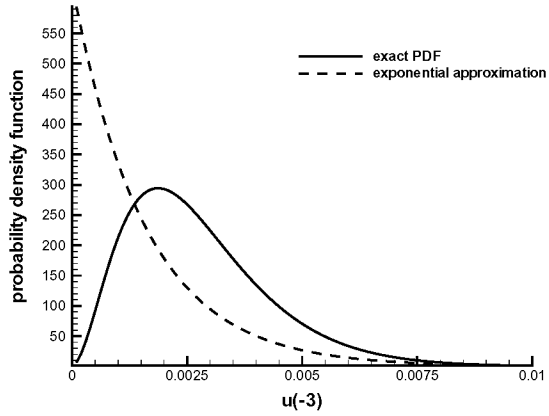


FIG. 11.2. *Comparison between approximate exponential and exact PDF at left-end boundary, computed using a random variable model for* μ.

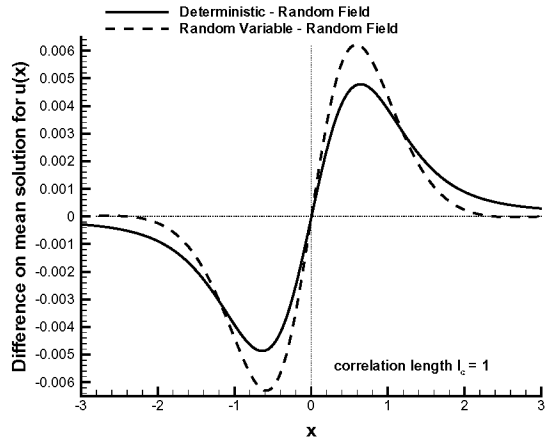


FIG. 11.3. *Difference between deterministic, mean random variable and mean random field solution (boundary condition Model (b), exponential auto-correlation with* $l_c = 1$*).*

12. Summary. The paper describes and highlights the need for more advanced uncertainty modeling. The different sources of uncertainty and their treatment are reviewed. The steady-state generalized Burgers equation (GBE) is solved using a variety of stochastic models. First a random variable model is used. The differential equation is solved deterministically and the inherent uncertainty associated with the viscosity is propagated through the closed-form solution. The exact solutions for the mean and standard deviation are compared with Monte Carlo simulations as well as with first- and second-order approximate results. The second-order results are found to be significantly more accurate.

Subsequently, the differential equation is solved stochastically using a random field description for the viscosity. The viscosity generally depends on the temperature distribution for which no information is available in the GBE. The random field emulates the model uncertainty caused by these missing variables. We computed results using both midpoint and locally averaged discretizations. It is shown how, for a given discretization level, both methods bracket the true variability. They provide practical upper and lower bounds to the variability. The impact of the auto-correlation model is studied as well.

The use of a random field model for the viscosity instead of a single random variable transforms the Burgers

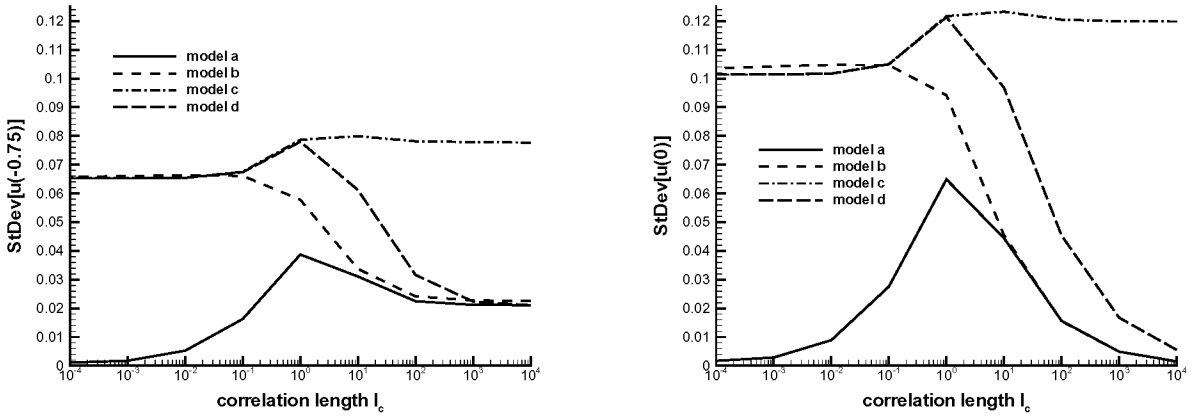


Figure a: $x = -0.75$ Figure b: $x = 0$

FIG. 11.4. *Impact of boundary condition uncertainty on* StDev$[u(x)]$ *for different correlation lengths* l_c.

equation into a stochastic differential equation. The presented results clearly reveal the random field effect: Var$[u(0)]$ depends on both the correlation length l_c and the auto-correlation model $R(\xi)$. It is intuitively clear that a zero variance at $x = 0$, as obtained from the random variable model, is unrealistic. The uncertainty caused by the missing variable – temperature – can be modeled realistically using a random field description of the viscosity.

In the last section, the effects of boundary condition uncertainty are studied. This uncertainty is caused by the truncation of the integration domain at the far field boundary during the numerical solution phase of the differential equations. In practice only limited information will be available regarding this type of model uncertainty. Therefore we compared various approximate models. It can be concluded that ignoring this boundary condition uncertainty dramatically underestimates the variance of the velocity $u(x)$ in the interior of the domain.

REFERENCES

[1] AIAA APPLIED AERODYNAMICS TECHNICAL COMMITTEE, *Drag Prediction Workshop, 19th APA Conference*, Anaheim, CA, 2001. http://www.aiaa.org/tc/apa/dragpredworkshop/dpw.html.

[2] *Guide for the verification and validation of computational fluid dynamics simulations*. AIAA Guide G-077-1998, 1998.

[3] D. A. ANDERSON, J. C. TANNEHILL, AND R. H. PLETCHER, *Computational Fluid Mechanics and Heat Transfer*, Series in computational methods in mechanics and thermal sciences, McGraw-Hill, 1997.

[4] A. H. S. ANG AND W. K. TANG, *Probability Concepts in Engineering Planning and Design, Volume I, Basic Principles*, John Wiley & Sons, 1975.

[5] Y. BEN-HAIM AND I. ELISHAKOFF, *Convex Models of Uncertainty in Applied Mechanics*, Elsevier Science, 1990.

[6] A. DER KIUREGHIAN, *Bayesian Analysis of Model Uncertainty in Structural Reliability*, in Third IFIP Working Conference on Reliability and Optimization of Strutural Systems, Berkeley, CA, 1990, Springer, pp. 211–221.

[7] A. DER KIUREGHIAN AND J.-B. KE, *The Stochastic Finite Element Method in Structural Reliability*, Probabilistic Engineering Mechanics, 3 (1988), pp. 83–91.

[8] A. DER KIUREGHIAN AND P.-L. LIU, *Structural Reliability under Incomplete Probability Information*, ASCE Journal of Engineering Mechanics, 112 (1986), pp. 85–104.

[9] G. A. FENTON, *Error Evaluation of Three Random-Field Generators*, ASCE Journal of Engineering Mechanics, 120 (1994), pp. 2478–2497.

[10] G. A. FENTON AND E. H. VANMARCKE, *Simulation of random fields via local average subdivision*, ASCE Journal of Engineering Mechanics, 116 (1990), pp. 1733–1749.

[11] S. FERSON, *Naive monte carlo methods yield dangerous underestimates of tail probabilities*, in High Consequence Operations Safety Symposium, J. A. Cooper, ed., Albuquerque, NM, 1994, Sandia National Laboratories.

[12] R. V. FIELD, JR. AND J. R. RED-HORSE, *A probabilistic approach to quantifying model uncertainty*, in 8^{th} International Conference on Structural Safety and Reliability, R. Corotis, G. Schuëller, and M. Shinozuka, eds., Newport Beach, CA, June 2001.

[13] A. HALDAR AND S. MAHADEVAN, *Probability, Reliability and Statistical Methods in Engineering Design*, John Wiley & Sons, 1999.

[14] J. S. HAMMOND, H. RAIFFA, AND R. L. KEENEY, *Smart Choices: A Practical Guide to Making Better Decisions*, Harvard Business School Publishing, 1999.

[15] E. R. HANSEN, *Global Optimization Using Interval Analysis*, Marcel Dekker, 1992.

[16] R. HOGG AND A. CRAIG, *Introduction to Mathematical Statistics*, Macmillan Publishing Co., fourth ed., 1978.

[17] L. HUYSE, *Free-form airfoil shape optimization under uncertainty using maximum expected value and second-order second-moment strategies*, tech. rep., ICASE Report 2001-18/NASA CR 2001-211020, 2001.

[18] M. G. KENDALL AND A. STUART, *The Advanced Theory of Statistics - Volume 2*, Hafner Publishing Company, New York, NY, 1961.

[19] R. S. LANGLEY, *Unified approach to probabilistic and possibilistic analysis of uncertain systems*, ASCE Journal of Engineering Mechanics, 126 (2000), pp. 1163–1172.

[20] M. A. LAWRENCE, *Basis random variables in finite element analysis*, International journal for Numerical Methods in Engineering, 24 (1987), pp. 1849–1863.

[21] C.-C. LI AND A. DER KIUREGHIAN, *Optimal Discretization of Random Fields*, ASCE Journal of Engineering Mechanics, 119 (1993), pp. 1136–1154.

[22] W. K. LIU, T. BELYTSCHKO, AND Y. J. LUA, *Probabilistic Finite Element Method*, in Probabilistic Structural Mechanics Handbook - Theory and Industrial Applications, C. Sundarajan, ed., London, UK, 1995, Chapman & Hall, pp. 70–105.

[23] W. K. LIU, T. BELYTSCHKO, AND A. MANI, *Random Field Finite Elements*, International Journal on Numerical Methods in Engineering, 23 (1986), pp. 1831–1845.

[24] H. MADSEN, S. KRENK, AND N. LIND, *Methods of Structural Safety*, Prentice-Hall, Englewood Cliffs, NJ, 1986.

[25] R. E. MELCHERS, *Structural Reliability Analysis and Prediction*, John Wiley & Sons, 1999.

[26] M. A. MEYER AND J. M. BOOKER, *Eliciting and Analyzing Expert Judgement*, U.S. Nuclear Regulatory Commission, Washington, D.C., 1990.

[27] T. NAGANUMA, G. DEODATIS, AND M. SHINOZUKA, *Arma model for two-dimensional processes*, ASCE Journal of Engineering Mechanics, 113 (1987), pp. 234–251.

[28] W. L. OBERKAMPF, J. C. HELTON, AND K. SENTZ, *Mathematical representation of uncertainty, aiaa-paper 2001-1645*, in 42^{nd} AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference and Exhibit, Seattle, WA, 2001. CD-ROM.

[29] A. PAPOULIS, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill Series in Electrical Engineering, McGraw-Hill, New York, NY, third ed., 1991.

[30] M. M. PUTKO, P. A. NEWMAN, A. C. TAYLOR III, AND L. L. GREEN, *Approach for Uncertainty Propagation and Robust Design in CFD Using Sensitivity Derivatives, AIAA Paper 2001-2528*, in AIAA 15^{th} Computational Fluid Dynamics Conference, Anaheim, CA, June 2001.

[31] S. S. RAO AND J. P. SAWYER, *Fuzzy finite element approach for the analysis of imprecisely defined systems*, AIAA Journal, 33 (1995), pp. 2364–2370.

[32] P. J. ROACHE, *Quantification of uncertainty in computational fluid dynamics*, Annual Review of Fluid Mechanics, 29 (1997), pp. 123–160.

[33] E. SAMARAS, M. SHINOZUKA, AND A. TSURUI, *Arma representation of random processes*, ASCE Journal of Engineering Mechanics, 111 (1985), pp. 449–461.

[34] G. I. SCHUËLLER, *A State-of-the-Art Report on Computational Stochastic Mechanics*, Probabilistic Engineering Mechanics, (1997), pp. 197–321.

[35] G. A. F. SEBER AND C. J. WILD, *Nonlinear Regression*, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, New York, NY, 1989.

[36] G. SHAFER, *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, NJ, 1976.

[37] M. SHINOZUKA AND G. DEODATIS, *Simulation of Stochastic Processes by Spectral Representation*, Applied Mechanics Reviews, 44 (1991), pp. 191–203.

[38] M. SHINOZUKA AND C.-M. JAN, *Digital Simulation of Random Processes and Its Applications*, Journal of Sound and Vibration, 25 (1972), pp. 111–128.

[39] P. D. SPANOS AND R. GHANEM, *Stochastic finite element expansion for random media*, ASCE Journal of Engineering Mechanics, 115 (1989), pp. 1035–1053.

[40] E. VANMARCKE, *Random Fields: Analysis and Synthesis*, The MIT Press, Cambridge, MA, 1983.

[41] R. W. WALTERS AND L. HUYSE, *Stochastic methods for fluid mechanics - an introduction*, Tech. Rep. in preparation, ICASE, NASA Langley Research Center, Hampton, VA, 2001.

[42] S. WOLFRAM, *The Mathematica Book*, Wolfram Media, Cambridge University Press, fourth ed., 1999.

[43] D. XIU AND G. E. KARNIADAKIS, *The wiener-askey polynomial chaos for stochastic differential equations*, Tech. Rep. 01-1, Center for Fluid Mechanics, Division of Applied Mathematics, Brown University, Providence. RI, 2001.

[44] F. YAMAZAKI AND M. SHINOZUKA, *Simulation of Stochastic Fields by Statistical Preconditioning*, ASCE Journal of Engineering Mechanics, 116 (1990), pp. 268–287.

[45] J. ZHANG AND B. ELLINGWOOD, *Orthogonal Series Expansion of Random Fields in Reliability Analysis*, ASCE Journal of Engineering Mechanics, 120 (1994), pp. 2660–2676.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	October 2001	Contractor Report

4. TITLE AND SUBTITLE
Random field solutions including boundary condition uncertainty for the steady-state generalized Burgers equation

5. FUNDING NUMBERS
C NAS1-97046
WU 505-90-52-01

6. AUTHOR(S)
Luc Huyse and Robert W. Walters

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
ICASE
Mail Stop 132C
NASA Langley Research Center
Hampton, VA 23681-2199

8. PERFORMING ORGANIZATION REPORT NUMBER
ICASE Report No. 2001-35

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
National Aeronautics and Space Administration
Langley Research Center
Hampton, VA 23681-2199

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
NASA/CR-2001-211239
ICASE Report No. 2001-35

11. SUPPLEMENTARY NOTES
Langley Technical Monitor: Dennis M. Bushnell
Final Report
To be submitted to the Journal of Fluid Mechanics

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Unclassified–Unlimited
Subject Category 64
Distribution: Nonstandard
Availability: NASA-CASI (301) 621-0390

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

CFD results are subject to considerable uncertainty associated with the operating conditions. Even when the operational uncertainty is omitted under very controlled circumstances during wind tunnel experiments, substantial disagreement between experimental and CFD results persists. This discrepancy must be attributed to model uncertainty. This report discusses the various sources of uncertainty. The need for advanced uncertainty modeling is illustrated by means of a computationally inexpensive 1-D Burgers equation model. We specifically address the uncertainty due to missing variables (inexact or incomplete differential equations). To this extent a random field model is used for the viscosity and the fundamental differences between the solutions of the stochastic differential equations and a simple random variable model is highlighted. The Burgers equation theoretically needs to be integrated over an infinite domain. In a deterministic approach, the integration domain is cut off at some far field boundary. This truncation effectively ignores all variability outside this far field boundary. We present a practical treatment for the uncertainty on the boundary conditions. The results indicate that ignoring the boundary condition uncertainty dramatically underestimates the variance of the velocity in the interior of the domain.

14. SUBJECT TERMS	15. NUMBER OF PAGES
model uncertainty, random field, uncertain boundary conditions	30
	16. PRICE CODE A03

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified		